AD/A-001 265

TESTING FOR A MONOTONE TREND IN A MODULATED RENEWAL PROCESS

Peter A. W. Lewis, et al

Naval Postgraduate School

Prepared for:

Office of Naval Research

December 1973

NPS-55Lw73121

NAVAL POSTGRADUATE SCHOOL

Monterey, California

TESTING FOR A MONOTONE TREND

IN A MODULATED RENEWAL PROCESS

P. A. W. Lewis

and

D. W. Robinson

December 1973

Prepared for:
Office of Naval Research, Arlington, Virginia 22217

24

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral Mason Freeman
Superintendent

Jack R. Borsting
Provost

The work reported herein was supported in part by the Foundation Research Program of the Naval Postgraduate School with funds provided by the Chief of Naval Research.

This report was prepared by:

PETER A. W. LEWIS, Professor
Department of Operations Research
and Administrative Sciences

DAVID W. ROBINSON
Department of Mathematics

Reviewed by:

DAVID A. SCHRADY, Chairman
Department of Operations Research
and Administrative Sciences

Released by:

DAVID B. HOISINGTON
Acting Dean of Research

ACCESSION for

NTIS White Section ☒
DDC Buff Section ☐
UNANNOUNCED ☐
JUSTIFICATION ....

....

BY ....
DISTRIBUTION/AVAILABILITY CODES

Dist.	AVAIL. and/or SPECIAL
A	

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD/A-001 265

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER NPS-55Lw73121	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Testing for a Monotone Trend in a Modulated Renewal Process		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Peter A. W. Lewis David W. Robinson		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N,RR 014-05-01, NR 042-284,PO-4-0025
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE December 1973 13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Chief of Naval Research Arlington, Virginia 22217		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Trend analysis
Modulated renewal process
Poisson process
Point processes
Log transforms

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In examining point processes which are overdispersed with respect to a Poisson process, there is a problem of discriminating between trends and the appearance in data of sequences of very long intervals. In this case the standard "robust" methods for trend analysis based on log transforms and regression techniques perform very poorly, and the standard exact test for a monotone trend derived for modulated Poisson processes is not robust with respect to its distribution theory when the underlying process is

(cont. next page)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

non-Poisson. However, experience with data and an examination of the departures from the Poisson distribution theory suggest a modification to the standard test for trend, both for modulated renewal and general point processes. The utility of the modified test statistic is verified by examining several sets of data, and simulation results are given for the distribution of the test statistic for several renewal processes.

Testing for a Monotone Trend in a Modulated Renewal Process

P. A. W. Lewis* and D. W. Robinson*

Abstract. In examining point processes which are overdispersed with respect to a Poisson process, there is a problem of discriminating between trends and the appearance in data of sequences of very long intervals. In this case the standard "robust" methods for trend analysis based on log transforms and regression techniques perform very poorly, and the standard exact test for a monotone trend derived for modulated Poisson processes is not robust with respect to its distribution theory when the underlying process is non-Poisson. However, experience with data and an examination of the departures from the Poisson distribution theory suggest a modification to the standard test for trend, both for modulated renewal and general point processes. The utility of the modified test statistic is verified by examining several sets of data, and simulation results are given for the distribution of the test statistic for several renewal processes.

1. Introduction. Stochastic point processes or series of events can be described either through the sequence of times to events $\{T_i\}$, or through the counting process $\{N_t\}$, where N_t is the number of events occurring in $(0,t]$. Trends on both serial number i and on time t are possible, but we only consider the time trends here, nor do we consider grouped data.

A fairly complete description of trend analysis for Poisson point processes is given in Cox and Lewis [4], Lewis [11], Lewis [10]and Brown [2]. In these works there is another minor difference which complicates matters; this is that observation may be for a fixed time interval $(0,t_0]$ or for a fixed number n of events. Fixed time observation is more common in practice but the fixed number case is easier to simulate, so we consider both, depending on convenience. Except for messy details the results are essentially the same.

We will also consider only the case of a simple monotone trend in time for the process, extending the Poisson theory to the case of more general point

*Naval Postgraduate School, Monterey, California. This research was supported by the Office of Naval Research through grant NR042-284.

processes. In the case of a non-homogeneous or modulated Poisson process a simple model [4, pp. 45] for the rate $\lambda(t)$,

$$(1) \qquad \lambda(t) = \exp\{\alpha+\beta t\} = \lambda\exp\{\beta t\}, \qquad t>0,\ \lambda>0,$$

leads to a uniformly most powerful conditional test for $\beta=0$ against $\beta\neq0$ based on the statistic

$$\sum_{i=1}^{N_{t_0}} T_i .$$

The conditioning is on N_{t_0}, the observed number of events in $(0,t_0]$, since N_{t_0} is a sufficient statistic for the nuisance parameter α for all β. Conditionally the statistic has mean $N_{t_0}/2$ and variance $N_{t_0}^2/12$, so the statistic

$$(2) \qquad U = \frac{\sum(T_i/t_0) - \frac{n}{2}}{(n/12)^{\frac{1}{2}}} ,$$

which converges rapidly to a unit normal variable under the null hypothesis, is used to test for $\beta=0$. The hypothesis is rejected for large or small values of U.

The test statistic U is computed in the SASE IV program for the analysis of point processes [13] and the program stops if $|U|>1.96$, since subsequent analysis in the program is for stationary processes. However, most users bypass this stop because it almost always occurs. This has led to the present work, the supposition being that the distribution theory of U is very sensitive to the Poisson hypothesis. Two sets of data which lead to this program stop are discussed in the next section. Then other possible test statistics are discussed Section 3, and the distribution of a statistic similar to $\sum T_i$ is examined for the special case of a Gamma renewal process. This leads to a simple modification of the test statistic to account for the overdispersion of the intervals between events relative to the exponential distribution.

In subsequent sections simulation results for the null distribution of the statistic are given for other renewal processes. Then the modification of the test which is required for general point processes is discussed. It is the simplicity of the extension in this general case which makes the test statistic attractive when compared to other possibilities. The problem of the power of different tests for trend has not been considered.

Finally we note that the situation we are interested in is that in which the point process is overdispersed with respect to the Poisson process. This will be defined to be the situation in which the index of dispersion for counts [4, pp. 71],

$$I = \lim_{t\to\infty} J(t) = \frac{\operatorname{var}(N_t)}{E(N_t)},$$

is greater than one, its value for the Poisson process. For the most part this corresponds to the marginal distribution of times between events having a coefficient of variation

$$C(x) = \frac{\sigma(x)}{E(x)}$$

greater than 1. This is always true for renewal processes, and for cluster processes (see [12] and [8]).

2. Data Analysis. Two sets of data are examined here and the results of tests for trend based on U are discussed.

Statistics for the first set are tabulated in Table 1. This set consists of 3 sequences of page exceptions in a multiprogrammed two-level memory computer with demand paging [14]. There is no particular compelling reason to expect a monotone trend in the data, except for an initial transient. This transient occurs because no page exception can occur until the memory is filled to the exception levels, which are 76, 197, and 512 in the three sequences examined. The transient is almost negligible at level 76, where the test based on U (column 4) rejects homogeneity at a 1% level. The rejection is stronger for the other levels, and at exception level 512 there is a very long transient and therefore inhomogeneity.

Note however that the intervals between events are very skewed with respect to the exponential distribution, the coefficients of variation given in column 5 being on the order of 3, compared to 1 for an exponentially distributed variate, and the coefficients of skewness γ_1 given in column 6 of Table 1 being greater than the value $\gamma_1=2$ for the exponential distribution.

An even more striking failure for the test occurs in the second set of data explored in Table 2. The events are occurrences of earthquakes with energies greater than 4.0 on the Richter scale in California and Nevada from 1932 to 1969. Six sections with equal numbers of events (except for the last) were analyzed and their statistics are given on the first six rows of Table 2. Columns 5 to 7 show that the intervals are very skewed, and the estimated serial correlation coefficients $\hat{\rho}_1$ in column 8 show the intervals to be correlated.

There is no particular reason to expect a monotone trend in this data, but $|U|$ is greater than 1.96 for all sections. The average of the U values is -0.72 and the estimate of the standard deviation of U for the sections (the sample standard deviation of the 6 U's) is shown in row 9, column 4 to be $\hat{\sigma}=7.82$. This is far in excess of the value of $\sigma=1$ for the U statistic under the hypothesis of a homogeneous Poisson process.

We will return to this data later on.

3. General remarks on the test statistic. Neither of the series considered above can be modelled as a renewal process since the estimated first serial correlation coefficients $\hat{\rho}_1$ are large. In fact the first set has been modelled as a univariate semi-Markov process by Lewis and Shedler [14] and the earthquake data is well known to be some kind of cluster process (Lewis,[12]; Vere-Jones [18]).

It is useful to consider renewal situations however, even if they occur rarely in practice, because of analytical possibilities. Cox [3] has extended the model (1) to modulated renewal processes by defining the intensity function $\lambda(t)$ as

(3) $$\lambda(t) = z\{u(t)\} \exp\{\alpha+\beta t\} ,$$

Table 1. Page exceptions in a multiprogrammed two-level memory computer with demand paging

Level (# pages)	N_{t_0}	t_0 (page references)	U	$\hat{C}(x)$	$\hat{\gamma}_1$	$\hat{\rho}_1$	$\frac{U}{\{\hat{C}(x)\}}$
76	1,807	8,802,464	-2.83	3.34	10.34	+0.188	-0.85
197	820	8,802,464	-8.67	3.27	7.14	+0.177	-2.60
512	517	8,802,464	-18.11	3.70	6.87	+0.130	-4.9C

Table 2. Earthquake Data - All earthquakes with energies greater than 4.0 in California and Nevada; 1932-1969

Section	N_{t_0}	t_0 (hours)	U	$\hat{C}(x)$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\rho}_1$	$\frac{U}{\{\hat{C}(x)\}}$
1	468	72,200	4.4	1.8	5.50	42.9	+0.49	2.44
2	468	58,921	-6.7	1.65	3.67	22.4	+0.16	-4.06
3	468	49,733	9.9	1.70	2.80	12.8	+0.22	5.82
4	468	29,403	2.1	1.70	3.30	17.5	+0.14	1.23
5	468	48,061	-11.7	1.50	2.40	9.8	+0.34	-7.80
6	431	79,686	-2.3	1.25	2.40	12.6	+0.12	-1.84
Average			-0.72	1.6	3.01	19.67	0.245	-.702
$\hat{S}_{\overline{x}}$			(3.19)	(0.81)	(0.68)	(4.99)	(0.059)	
$\hat{\sigma}$			7.82	0.197	1.67	12.22	(0.144)	
TOTAL Record	2771	338,004	-0.527	1.63	-	-	-	-0.323

where $z(\cdot)$ is the hazard [4, pp. 135] or hazard rate in the terminology of some workers in reliability theory. However, although a complete likelihood can be set up [3] it has not been possible to derive any explicit tests for $\beta=0$ from it.

We therefore continue to examine modifications of the U statistic. For convenience, however, we consider the case of observation for a fixed number of events n. There are several reasons for this:

(i) The fixed number case is much simpler to simulate and statistical differences between the two situations will be minor, especially for large samples.

(ii) The sufficient statistic for α in the model (1) for a Poisson process is $Y_{1n} = \sum_{i=1}^{n} X_i$, where X_i are the times between events and the test statistic [4, p. 52] is

(4) $$Y_{2n} = \sum_{i=1}^{n} S_i$$

(5) $$= \sum_{i=1}^{n} (n+1-i)X_i$$

Although this statistic can be considered conditionally on Y_{1n}, it follows from well known characterizing results for exponential and Gamma distributed variates (see Lukacs and Laha [15]) that this is equivalent to considering the test statistic

(6) $$Y_n = \frac{Y_{2n}}{Y_{1n}} = \frac{\sum_{i=1}^{n} S_i}{\sum_{i=1}^{n} X_i} .$$

Moreover for any renewal model with intensity function (3) this statistic will be free of the nuisance parameter α for any β, as can be easily shown. This is an important simplification.

(iii) Analytical results for the fixed number case are simpler to obtain than those for the fixed time case. Moreover (6) suggests several other possibilities. From the form (5) for the numerator it can be seen that it is like an empirical serial correlation between the natural numbers and the serially ordered times between events X_i. This is the form of several standard tests for trend [7, Ch. 45]. A possibility would be to replace the X_i's by exponential scores and correlate the serially ordered scores with the index numbers i. Permutation tests of this sort have been discussed by Guillier [6]; we do not pursue them here because they depend on the independence assumption in the renewal hypothesis and we wish to consider more general point processes with dependent times-between-events.

Two other possible tests for trend are noted here.

One is based on log transformations of the data and standard regression techniques, but as noted in Cox and Lewis [4, pp. 41] these methods are likely to have poor relative power for intervals X_i which are more dispersed than exponential variates. (For fairly regular processes they are likely to be the favored procedures.)

The second possibility arises from an analogy between Y_n and goodness of fit tests. Define

(7) $$\xi_{n,i} = \frac{S_i}{\sum_{j=1}^{n} X_j} \qquad i=1, \ldots, (n-1).$$

Then if $\hat{F}_n(y)$ denotes [17] the empirical cumulative distribution function for $\xi_{n,i}$, i=1, .., (n-1), we have

(8) $$\int_0^1 \{\hat{F}_n(u) - u\}du = (n+1) - Y_n .$$

Thus Y_n is essentially a one sided Cramér-von Mises statistic and other norms could be tried to measure the deviation of $F_n(u)$ from the function u between 0 and 1.

Because the statistic Y_n and tests for trend based on it can be extended to non-renewal processes, we consider its distribution first for Gamma renewal processes, then for several other renewal processes and then for cluster processes.

<u>4. Testing in modulated Gamma renewal processes</u>. The Gamma renewal process has independently distributed intervals with probability density function [4, pp. 136]

$$(9) \qquad f_X(x) = \left(\frac{k}{\mu}\right)^k \frac{x^{k-1}e^{-kx/\mu}}{\Gamma(k)} \qquad x>0,\ k>0,$$

where $\Gamma(k)$ is the complete Gamma function. For $k=1$ we have an exponentially distributed variate, and for $k=\frac{1}{2}$ the square of a normal random variable. We will be concerned with the case $k\leq 1$. We also have

$$(10) \qquad E(X) = \mu;\ \text{var}\ (X) = \frac{\mu^2}{k};\ C(X) = \frac{1}{\sqrt{k}}\ .$$

Consider now the distribution of Y_n given by (6), which we write for convenience as

$$(11) \qquad Y_n = \frac{\sum_{i=1}^{n}(n+1-i)X_i/n}{\sum_{i=1}^{n}X_i/n} = \frac{Y'_{2n}}{Y'_{1n}}\ .$$

The moments of the numerator and denominator are

$$(12) \qquad E\left(Y'_{1n}\right) = \mu, \quad \text{var}\ \left(Y'_{1n}\right) = \sigma^2/n,$$

$$(13) \qquad E\left(Y'_{2n}\right) = (n+1)\mu/2, \quad \text{var}\ \left(Y'_{2n}\right) = (n+1)(2n+1)\sigma^2/(6n).$$

Now it is a characterizing property of Gamma distributed variates [15, p. 58] that the expected value of ratios of linear functions of the Gamma variates such as those appearing in (11) is the expected value of the ratio of the expectations. Thus we have, for Gamma renewal processes,

(14) $$E(Y_n) = (n+1)/2;$$

(15) $$\operatorname{var}(Y_n) = \frac{(n-1)}{12}\frac{(n+1)}{(kn+1)} = \frac{(n-1)}{12}\frac{(n+1)}{(n/C^2(x)+1)};$$

(16) $$\operatorname{var}(Y_n) \sim \frac{n-1}{12} C^2(x).$$

Since $C^2(x)$ equals one for a Poisson process ($k=1$), this checks with results for the statistic U given in (2).

Note further that

$$Y_n' = Y_n - \frac{n+1}{2} = \frac{\sum_{i=1}^{n} X_i\left(\frac{n+1}{2n} - \frac{i}{n}\right)}{\sum_{i=1}^{n} X_i/n}$$

(17) $$= \frac{\sum_{i=1}^{\left|\frac{n}{2}\right|}\left(X_i - X_{n+1-i}\right)\left(\frac{n+1}{n} - \frac{2i}{n}\right)}{\sum_{i=1}^{n} X_i/n}$$

(18) $$= \frac{\sum_{i=1}^{\left|\frac{n}{2}\right|} X_i' a_i}{\sum_{i=1}^{n} X_i/n},$$

where $\left|\frac{n}{2}\right|$ is the greatest integer less than or equal to $n/2$, $X_i' = X_i - X_{n+1-i}$ is a symmetric random variable and a_i is an odd sequence.

Using (18) we can show the following results:

(i) The centered statistic Y_n' has odd moments which are all zero. This follows because the numerator in (18) is a sum of independent symmetric random variables and is therefore [5, Lemma 2, p. 149] itself symmetric. This implies that the odd moments of the numerator (including the first) are zero and by the Lukacs and Laha result cited above, so are those of Y_n'.

Thus Y_n' is a symmetric random variable.

(ii) The numerator in (18) divided by $(n)^{\frac{1}{2}}$ is asymptotically normal. Moreover since the denominator converges with probability one to μ, which is non-zero, results from Billingsley [1, Corollary 2, p. 31] show that the reciprocal of the denominator converges with probability one to $1/\mu$. Slutsky's Theorem (see Billingsley [1]) then says that

$$(19) \qquad \frac{\left(\frac{12}{n-1}\right)^{\frac{1}{2}} Y_n'}{C(X)} \overset{L}{\to} N(0,1).$$

(iii) Convergence to the normal distribution is likely to be very rapid because of the symmetry of the distribution of Y_n'.

To examine the small sample distribution of Y_n for the Gamma renewal case an extensive simulation was undertaken. Detailed results are given in Robinson [16]. The results are illustrated in Table 3, which is extracted from Robinson [16].

The simulations involved 100,000 replications using the random number generator LLRANDOM (Learmonth and Lewis [9]) and a Gamma random number generator developed by Robinson [16]. The computations were checked by comparing the theoretical results for the mean and variance of the statistics with the simulated mean and variance.

Only the case k=0.1 $(C^2(X)=10)$ is given in Table 3 because this was the most extreme case simulated and has the greatest departure from normality and the slowest convergence to the asymptotic normal form. Simulated quantiles of Y_n, normalized by subtracting the mean (14) and dividing by the square root of the variance (15) (these are listed in the last two rows of the table) are shown in Table 3. Because of the symmetry of the distribution, only the lower quantiles corresponding to levels α=0.001, 0.002, 0.005, 0.010, 0.020, 0.025, 0.050, 0.100, 0.200, 0.300, 0.400, 0.500 are given. They are actually the average of the simulated upper and lower quantiles and have a standard deviation of approximately 0.001.

The distribution can be seen to be a little more peaked than a normal distribution, with shorter tails, but even by n=50 a normal approximation to the

Table 3. Simulation results for the statistic Y_n for Gamma distributed intervals with k=0.10 under the null hypothesis of no trend (β=0). Quantiles of Y_n are normalized by subtracting $E(Y_J)$ and dividing by $\sigma(Y_J)$.

α	n=10	n=30	n=50	n=100	Normal quantile
0.001	-2.202	-2.740	-2.915	-3.001	-3.090
0.002	-2.191	-2.607	-2.750	-2.812	-2.878
0.005	-2.148	-2.460	-2.500	-2.545	-2.576
0.010	-2.078	-2.231	-2.290	-2.313	-2.326
0.020	-1.944	-2.014	-2.049	-2.054	-2.054
0.025	-1.875	-1.935	-1.960	-1.965	-1.960
0.050	-1.654	-1.665	-1.665	-1.656	-1.645
0.100	-1.343	-1.320	-1.307	-1.297	-1.282
0.200	-0.924	-0.881	-0.871	-0.856	-0.842
0.300	-0.591	-0.554	-0.549	-0.537	-0.524
0.400	-0.279	-0.272	-0.265	-0.261	-0.253
0.500	-0.001	-0.005	0.002	-0.003	0.000
$E(Y_n)$	5.5	15.5	25.5	50.5	
$\sigma(Y_n)$	2.031	4.328	5.891	8.703	

distribution of Y_n is adequate for purposes of hypothesis testing.

The proposal for testing a monotone trend in a Gamma renewal process derived from these results is to estimate the coefficient of variation from the data and test for $\beta=0$ using

$$\frac{\left(\frac{12}{n-1}\right)^{\frac{1}{2}} Y_n'}{\{\hat{C}(X)\}}$$

and assuming that its distribution is that of a unit normal distribution. This essentially uses the Poisson test statistic divided by $\hat{C}(X)$. This modified statistic is given in the last columns of Tables 1 and 2. The test results are more in line with expectations, but still do not reflect inflation of the variance of U because of correlation between intervals between events. This is discussed in Section 6.

5. Distributional results for other renewal cases. The result (14) holds for any stationary sequence $X_1, \ldots, X_n$, including a renewal (i.i.d.) sequence. This is because

$$E\left(\frac{X_1 + \ldots + X_n}{X_1 + \ldots + X_n}\right) = nE\left(\frac{X_1}{X_1 + \ldots + X_n}\right) = 1,$$

or $E\{X_i/(X_1 + \ldots + X_n)\} = \frac{1}{n}$ for $i=1, \ldots, n$. Taking expectations in (6) and using the form (5) for Y_{2n} yields

$$E(Y_{2n}) = \frac{n+1}{2} .$$

This result merely says that Y_n, which is a normalized centroid of times to events in an interval stationary point process, always has the expected value $(n+1)/2$.

Thus the centering in (17) is correct for all sequences and we discuss Y_n' from here on.

Another useful result is that Y_n' is a symmetric random variable for any renewal sequence. To see this note that $-Y_n'$ can be written exactly in the form (18) with $X_i' = X_{n+1-i} - X_i$, but since these are symmetric random variables and the X_i's are independent, the functional form for $-Y_n'$ is exactly the same as that for Y_n'. Thus they have the same distribution and thus Y_n' is symmetrical random variable. All odd moments are thus zero. In addition by arguments of the previous section, Y_n' is asymptotically normal with variance (16) if $\text{var}(X)<\infty$ for any renewal process.

To explore the small sample distribution of Y_n further for renewal processes using simulation we chose two other density functions for the intervals.

The first is the Weibull density function

$$(20) \qquad f_X(x) = k\beta^k x^{k-1} \exp(-\beta^k x^k) \qquad \beta>0,\ k>0,\ x\geq 0$$

which reduces to the exponential for k=1. In the simulation the parameters were chosen so that the means and coefficients of variations of the intervals X were the same as for the Gamma cases.

The second density function chosen was the log-normal density, again with parameters chosen to match the means and coefficients of variations in the Gamma cases. Note that both these densities are, for given coefficient of variation, more skewed than the Gamma density, the log-normal more so than the Weibull. In addition both have hazard functions which approach zero as $x\to\infty$, in contrast to the Gamma density which has an exponential tail.

It is possible to compute $\text{var}(Y_n')$ for finite n in both these cases, but the results are messy. In general the variances are smaller than for the Gamma case; simulation results give, when $C^2(X) = 10.0$ and $n = 50$, values of 5.891, 5.182 and 4.355 for the Gamma, Weibull and log-normal cases respectively.

Only the worst case of the simulations for the Weibull and log-normal intervals, i.e., those matching the Gamma case with $C^2(X) = 10.0$ are given, in Table 4 and 5 respectively. Again 100,000 replications were used.

The normalized quantiles show distributions for Y_n' at n = 10, 30, 50, 100 for both densities and, in addition, for n = 200 for the log-normal case. In

Table 4. Simulation results for the statistic Y_n for Weibull distributed intervals with $C^2(x) = 10.0$ under the null hypothesis of no trend ($\beta = 0$). Quantiles of Y_n are normalized by subtracting $\tilde{E}(Y_n)$ and dividing by $\tilde{\sigma}(Y_n)$.

α	n = 10	n = 30	n = 50	n = 100	Normal Quantile
.001	-2.533	-2.922	-3.067	-3.214	-3.090
.002	-2.473	-2.772	-2.845	-2.973	-2.878
.005	-2.343	-2.521	-2.570	-2.635	-2.576
.010	-2.188	-2.301	-2.326	-2.373	-2.326
.020	-1.987	-2.042	-2.052	-2.069	-2.054
.025	-1.920	-1.954	-1.960	-1.971	-1.960
.050	-1.659	-1.652	-1.644	-1.641	-1.645
.100	-1.324	-1.294	-1.280	-1.272	-1.282
.200	-0.883	-0.850	-0.845	-0.831	-0.842
.300	-0.557	-0.531	-0.528	-0.516	-0.524
.400	-0.271	-0.255	-0.259	-0.249	-0.253
.500	-0.002	-0.000	0.000	-0.002	0.000
$\tilde{E}\, Y_n$	5.500	15.490	25.527	50.495	
$\tilde{\sigma}(Y_n)$	1.678	3.703	5.182	7.953	
$\tilde{\gamma}_1(Y_n)$	-0.002	-0.001	-0.001	-0.001	
$\tilde{\gamma}_2(Y_n)$	2.52	2.86	2.96	3.14	

Table 5. Simulation results for the statistic Y_n for log-normal distributed intervals with $C^2(x) = 10.0$ under the null hypothesis of no trend (β=0). Quantiles of Y_n are normalized by subtracting $\tilde{E}(Y_n)$ and dividing by $\tilde{\sigma}(Y_n)$.

α	n = 10	n = 30	n = 50	n = 100	n = 200	Normal Quantile
.001	-2.941	-3.342	-3.452	-3.692	-3.831	-3.090
.002	-2.805	-3.084	-3.167	-3.361	-3.411	-2.878
.005	-2.550	-2.725	-2.775	-2.845	-2.868	-2.576
.010	-2.325	-2.434	-2.445	-2.471	-2.472	-2.326
.020	-2.073	-2.104	-2.098	-2.106	-2.094	-2.054
.025	-1.975	-1.997	-1.991	-1.988	-1.978	-1.960
.050	-1.656	-1.638	-1.629	-1.621	-1.606	-1.645
.100	-1.289	-1.252	-1.246	-1.230	-1.224	-1.282
.200	-0.843	-0.813	-0.804	-0.791	-0.786	-0.842
.300	-0.524	-0.502	-0.495	-0.487	-0.484	-0.524
.400	-0.255	-0.241	-0.236	-0.236	-0.231	-0.253
.500	0.001	0.000	0.003	-0.003	0.003	0.000
$\tilde{E}(Y_n)$	5.501	15.484	25.518	50.492	100.463	
$\tilde{\sigma}(Y_n)$	1.365	3.059	4.355	6.889	10.699	
$\tilde{\gamma}_1(Y_n)$	-0.004	0.004	-0.015	-0.002	-0.001	
$\tilde{\gamma}_2(Y_n)$	2.89	3.35	3.51	3.87	4.11	

both cases the distributions have heavier tails than in the Gamma case, and estimated kurtoses γ_2 greater than one. The convergence to the asymptotic normal distribution is particularly slow for the log-normal case, but in no case is the normal approximation too far off at the quantiles corresponding to the usual significance levels used in hypothesis testing. Actually division of the quantiles by $C(X)[(n-1)/12]^{1/2}$ from (16) rather than by the true standard deviation of Y_n' provides a better normal approximation than does division of the quantiles by the true Var (Y_n').

Convergence is of course faster and the normal approximation better for the cases not shown here, i.e. for intervals with coefficients of variation approaching the value one of the exponential distribution. Note that $C^2(X) =$ 10.0 approximates the values found for the computer data of Table 1.

6. Distributional results for general point processes. The finding from the previous sections was that for renewal sequences the null hypothesis variance of Y_n' is inflated by approximately $C^2(X)$ over its value for a Poisson process. The approximation is exact for large n.

However, in both examples cited in Section 2 the intervals between events X_i are correlated (see the values $\hat{\rho}_1$ in Tables 1 and 2). It turns out that for a simple statistic such as Y_n' fairly broad results can be obtained for general point processes, the modification to the variance of Y_n' again being simple to compute from the data. Thus a rough test of trend can be performed.

Details of the derivation will be given elsewhere. For a broad class of situations Y_n' is asymptotically normally distributed with variance

$$(21) \qquad \text{var}\left(Y_n'\right) \cdot \frac{(n-1)}{12}\{\pi C^2(X) f_+(0+)\},$$

where $f_+(0+)$ is the initial point on the spectrum of the intervals $\{X_i\}$ of the process. Since $f_+(0+)$ is related to the initial point of the spectrum of counts, $g_+(0+)$, and the asymptotic slope, $V'(\infty)$, of the variance time curve, var $\{N_t\}$, of the point process by the relationship [4, p. 78]

$$(22) \qquad V'(\infty) = \pi g_+(0+) = \frac{\pi C^2(X)}{E(X)} f_+(0+) ,$$

we can write (21) as

$$\text{(23)} \qquad \operatorname{var}\left(Y_n'\right) \sim \frac{(n-1)}{12}\{E(X)V'(\infty)\}.$$

The quantity $V'(\infty)$ is simple to estimate from the data [4, pp. 115-120], thereby providing an easy modification for the test statistic Y_n'.

For a renewal process, $f_+(0+) = 1/\pi$, and (21) reduces to (16). Poisson cluster processes [12, 18] have been used to model the earthquake data of Section 2. If the length of the cluster in the cluster process is denoted by S, we have

$$\text{(24)} \qquad \operatorname{var}\left(Y_n'\right) \sim \frac{(n-1)}{12} E(S+1)\{1 + C^2(S+1)\},$$

where $C^2(S+1)$ is the coefficient of variation squared of S+1. When there is no cluster, i.e. S=0 with probability 1, the result (24) reduces to that for the Poisson process.

For the earthquake data, which has long and very variable clusters, the multiplier of (n-1)/12 in (24) has an estimated value of approximately 49.0. Dividing the U values given in column 4 of Table 2 by $(49)^{\frac{1}{2}}=7.0$, we obtain a test statistic which accepts the hypothesis of no trend in all 6 sections of the data.

<u>7. Conclusions and further work</u>. The recommendation put forward in this paper is to test for trend in a point process using the U statistic (2) divided by the estimated coefficient of variation $\hat{C}(X)$ in a renewal process, or an estimate of $\{E(X)V'(\infty)\}^{\frac{1}{2}}$ in (23) for a general point process.

The test is not proposed as being in any sense optimal, but because it can be used without detailed knowledge of the structure of the process it is very functional. It would be nearly optimal if the point process were close to a Poisson process.

The power of the test needs to be investigated so that its utility can be assessed relative to other tests, especially for processes which are highly overdispersed relative to the Poisson process. Point processes of that type occur in many applications.

Other tests to be considered could be standard regression tests after a log transform or scoring of the intervals in the data; rank correlation tests using, perhaps, exponential scores for the intervals, and other functionals than that given in (8) for measuring the "distance" of $F_n(u)$ from u (see [4, Ch. 6]). There are other possibilities explored in a recent thesis by Guillier [6].

REFERENCES

[1] P. BILLINGSLEY, Convergence of Probability Measures, Wiley, New York, 1962.

[2] M. BROWN, Statistical Analysis of Non-Homogeneous Poisson Processes, in Stochastic Point Processes, P.A.W. Lewis (ed.), Wiley, New York, 1972, pp. 67-89.

[3] D.R. COX, The Statistical Analysis of Dependencies in Point Processes, in Stochastic Point Processes, P.A.W. Lewis (ed.), Wiley, New York, 1972, pp. 55-66.

[4] D.R. COX and P.A.W. LEWIS, The Statistical Analysis of Series of Events, Methuen, London and Barnes and Noble, New York, 1966.

[5] W. FELLER, An Introduction to Probability Theory and its Applications, Volume 2, Second Edition, Wiley, New York, 1971.

[6] C.L. GUILLIER, Asymptotic Relative Efficiencies of Rank Tests for Trend Alternatives, PhD Thesis, University of California, Berkeley, June 1972.

[7] M.G. KENDALL and A. STUART, The Advanced Theory of Statistics, Vol. 3, Hafner, New York, 1966.

[8] A.J. LAWRANCE, Some Models for Stationary Series of Univariate Events, in Stochastic Point Processes, P.A.W. Lewis (ed.), Wiley, New York, 1972, pp. 199-256.

[9] G.P. LEARMONTH and P.A.W. LEWIS, Naval Postgraduate School Random Number Generator Package LLRANDOM, Report NPS 55LW 73061A, Naval Postgraduate School, Monterey, California, 1973.

[10] P.A.W. LEWIS, Recent Results in the Statistical Analysis of Univariate Point Processes, in Stochastic Point Processes, P.A.W. Lewis (ed.), Wiley, New York, 1972, pp. 1-54.

[11] P.A.W. LEWIS, Remarks on the Theory, Computation and Application of the Spectral Analysis of Series of Events, J. Sound Vib., 12(1970), pp. 353-375.

[12] P.A.W. LEWIS, Asymptotic Properties and Equilibrium Conditions for Branching Poisson Processes, J. Appl. Prob., 6(1969),pp. 355-371.

[13] P.A.W. LEWIS, A.M. KATCHER and A.H. WEISS, SASE IV - An Improved Program for the Statistical Analysis of Series of Events, IBM Research Report, RC 2365, 1969.

[14] P.A.W. LEWIS and G.S. SHEDLER, Empirically Derived Micromodels for Sequences of Page Exceptions, IBM Jour. Res. Dev., 17(1973), pp. 86-100.

[15] E. LUKACS, and R.G. LAHA, Applications of Characteristic Functions, Griffin, London, 1964.

[16] D.W. ROBINSON, (1973), The Generation of Gamma Distributed Variates and an Investigation of a Trend Test for the Gamma Renewal Process, Naval Postgraduate School Thesis, June 1973.

[17] G.R. SHORACK, Convergence of a Renewal Process, Abstract 72t-97 in Bulletin of Institute of Mathematical Statistics, 1, 5(1972), pp. 255-256.

[18] D. VERE-JONES, Stochastic Models for Earthquake Occurrence, J.R. Statist. Soc. B, 32(1970), pp. 1-62.